Published in final edited form as:

Stat Methods Med Res. 2019 January ; 28(1): 70–83. doi:10.1177/0962280217715051.

Bayesian analysis of longitudinal dyadic data with informative missing data using a dyadic shared-parameter model

Jaeil Ahn^{1,*}, Satoshi Morita², Wenyi Wang³, and Ying Yuan⁴

¹Department of Biostatistics, Georgetown University, Washington, DC 20057

²Department of Biomedical Statistics and Bioinformatics, Kyoto University, Japan

³Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030

⁴Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030

Summary

Analyzing longitudinal dyadic data is a challenging task due to the complicated correlations from repeated measurements and within-dyad interdependence, as well as potentially informative (or non-ignorable) missing data. We propose a dyadic shared-parameter model to analyze longitudinal dyadic data with ordinal outcomes and informative intermittent missing data and dropouts. We model the longitudinal measurement process using a proportional odds model, which accommodates the within-dyad interdependence using the concept of the actor-partner interdependence effects, as well as dyad-specific random effects. We model informative dropouts and intermittent missing data using a transition model, which shares the same set of random effects as the longitudinal measurement model. We evaluate the performance of the proposed method through extensive simulation studies. As our approach relies on some untestable assumptions on the missing data mechanism, we perform sensitivity analyses to evaluate how the analysis results change when the missing data mechanism is misspecified. We demonstrate our method using a longitudinal dyadic study of metastatic breast cancer.

Keywords

dyadic; non-ignorable missingness; longitudinal study; shared-parameter; intermittent missing; sensitivity analysis

1 Introduction

Dyadic data are collected from pairs of individuals (i.e., dyad), such as married couples, siblings, co-workers, and twins. The two members of a dyad are emotionally or behaviorally related because of sharing common genetic or/and environmental qualities. For instance, measures of marital happiness among married couples are likely to be interdependent, and

* ja1030@georgetown.edu.

one spouse's happiness strongly influences the happiness of the other (Gottman et al., 2002; Cook et al., 1995). In econometric and psychometric literature, peer effects, i.e., generalized form of dyadic interdependence, are frequently assumed in this context (Gardner and Steinberg, 2005; Carrell, Fullerton, and West, 2009). In this context, dyadic studies have put great emphasis on understanding the interdependence between the two members within a dyad.

This article is motivated by a longitudinal dyadic study that investigated depression in patients with metastatic breast cancer (MBC) and their spouses (Badr, Carmack, Karshy et al., 2010). For MBC patients, the spouse's support is one of the most important resources for coping with stress and depression, and a satisfactory spousal relationship shows association with lower levels of distress (Fang, Manne, and Pape, 2001; Carmack, Taylor et al., 2008). One of the primary objectives of the study was to characterize the level of depression among MBC patients and their spouses in relation to chronic psychological stress or physical pain, and to understand the dyadic interplay between MBC patients and their spouses. As such, better understanding the spousal interaction is critical for the future development of couple-based interventions that help MBC patients and their partners to manage depression.

The study initially proposed to recruit 334 participants (167 couples) at baseline, based on the identification of female patients who were initiating treatment for metastatic breast cancer using medical chart review during routine clinic visits. Project staff approached patients who had initiated treatment for metastatic breast cancer and asked them to sign the consent form for study participation. Patients who agreed to participate in the study were asked to complete a study questionnaire. With the patient's consent, spousal information was also gathered. The same questionnaire was administered again at 3 and 6 months after the baseline measurement. The questionnaire included questions regarding psychological functioning and the depression levels of the patient and spouse. Based on the Center for Epidemiology Studies Depression Scale (CES-D), depression is categorized into three levels: no depression, mild depression, and major depression, for the ease of interpretation (Zich et al., 1990). During the collection of data for the study, numerous dropouts occurred and data were missed intermittently (i.e., patients returned to the clinic for measurements after missing some visits before the end of the study). At baseline, data assessing the severity of depression were collected for 167 couples. After 3 months, 108 couples remained in the study; and at 6 months, only 91 couples completed the study. Sixteen patients and 15 spouses who had not attended the assessment at 3 months came back to the clinic to provide data at 6 months. The numbers of observations at each visit are summarized in Table 1. The investigator worried that the dropouts and intermittent data missingness might have been related to the change of the level of depression, and that the participants with deteriorating depression might have been more likely to miss a visit or drop out of the study. In other words, the missing data might be informative or non-ignorable. In addition, it is not clear whether the missing data mechanism might differ between patients and their spouses because the patients are directly impacted by the disease.

The analysis of such longitudinal dyadic data is challenging because of the requirement to account for both the within-dyad correlations and within-subject correlations (due to repeated measurements), as well as potentially non-ignorable dropout and intermittent

missing data. Statistical analysis of longitudinal data with non-ignorable missing data has been extensively investigated in the statistical literature, including selection models (Wu and Carroll, 1988; Diggle and Kenward, 1994; Follman and Wu, 1995), pattern-mixture models (Wu and Bailey, 1989; Little, 1993, 1994; Hogan and Laird, 1997; Michiels et al., 1999; Roy and Daniels, 2008), shared parameter models (SPM), which can be viewed as a special case of mixed-effect selection models (Follman and Wu, 1995), and the mixed-effects hybrid model (Yuan and Little, 2009; Ahn et al., 2013). Comprehensive reviews on methods to handle non-ignorable missing data in longitudinal studies are provided by Little (1995, 2008), Verbeke and Molenberghs (2000), Ibrahim and Molenberghs (2009), Hogan, Roy, and Korkontzelou (2004), and Daniels and Hogan (2008). Most of these existing methods focus on conventional longitudinal data, which is characterized by mutual independence between study participants. Thus, these methods cannot be directly applied to longitudinal dyadic data, which is characterized by correlated pairs of study participants.

Several approaches have been developed to handle longitudinal dyadic data with non-ignorable dropouts. Zhang and Yuan (2012) proposed a selection-model-based approach to handle non-ignorable dropouts. Ahn et al. (2013) proposed a latent mixed-effect hybrid model to address non-ignorable dropouts. These two methods assume that the outcomes are normally distributed and mainly focus on dropouts without considering intermittent missing data. In this paper, we propose a dyadic shared-parameter model (DSPM) using Bayesian inference to analyze the longitudinal dyadic data, where the outcomes are ordinal and subject to potentially non-ignorable intermittent data missingness. Such an approach has practical importance for two primary reasons: (1) a large percentage of psychosocial measures are ordinal in nature and are scored on the Likert scale; and (2) intermittent missing data commonly occur in practice. In the proposed DSPM, we model the outcome process using a cumulative proportional odds model, and model the missing data process using a transition model. We then link these two processes by letting them “share” a set of common random effects. Under the Bayesian framework, we develop the Markov chain Monte Carlo (MCMC) algorithm to fit the DSPM and make inferences. As with all methods dealing with non-ignorable missing data, our approach makes untestable assumptions on the missing data mechanism. We conduct extensive sensitivity analyses to evaluate the performance of the proposed approach when the missing data mechanism is misspecified.

In Section 2, we describe our DSPM and its three ingredients in detail. In Section 3, we introduce prior and posterior distributions. In Section 4, we compare the performance of our model with those of alternative approaches through various simulated scenarios. The application to the MBC data follows in Section 5. We conclude with a brief discussion in Section 6.

2 Probability model

2.1 Notation

Suppose a longitudinal dyadic study is designed to collect K repeated measurements of Y , paired ordinal outcomes, for each of n dyads. We assume that Y has L categories, i.e., $Y \in (1, 2, \dots, L)$. Let Y_{ijk} denote the ordinal outcome measured at the k th time point for the j th member of dyad i , with $\mathbf{Y}_{ij} = (Y_{ij1}, \dots, Y_{ijK})^\top$ and $\mathbf{Y}_i = (\mathbf{Y}_{i1}, \mathbf{Y}_{i2})^\top$. For example, in our

study, $K = 3$ measurements, and Y_{i1} and Y_{i2} are the response vectors for the i th dyad (MBC patient and spouse), respectively.

We here distinguish two types of missing data patterns. We call that 1) Y_{ijk} is intermittently missing if Y_{ijk} is missing but at least one of the later measurements ($Y_{ij,k+1}, \dots, Y_{ijK}$) is observed, that is, the participant missed the k th measurement but reappeared for later measurements; and 2) Y_{ijk} is missing due to dropout if Y_{ijk} and all subsequent measurements ($Y_{ij,k+1}, \dots, Y_{ijK}$) are missing, that is, the participant missed the k th measurement and never came back for later measurements. Although it could be possible that some dropouts might have reappeared subsequently if the study had continued beyond K measurements, we cannot identify these subjects based on the observed data and thus we will not consider that possibility here. To this end, we define the missing data indicator D_{ijk} as

$$D_{ijk} = \begin{cases} 0 & \text{if } Y_{ijk} \text{ is observed,} \\ \mathcal{I} & \text{if } Y_{ijk} \text{ is intermittently missing,} \\ \mathcal{D} & \text{if } Y_{ijk} \text{ is missing due to early drop out,} \end{cases}$$

with $D_{ij} = (D_{ij1}, \dots, D_{ijK})^T$ and $D_i = (D_{i1}, D_{i2})^T$. Let $X_i = (X_{i1}, X_{i2})^T$ denote the covariate matrix for the i th dyad where X_{ij} is a vector of length p , containing the p variable measurements for the j th member of the i th dyad.

2.2 Outcome model

We model the outcome (Y_{i1}, Y_{i2}) for the i th dyad using the cumulative proportional odds model as follows,

$$\text{logit}\{\Pr(Y_{i1k} \leq l | X_i, Z_{i1}, b_{i1})\} = \beta_{01l} - b_{i1}^T Z_{i1} - X_{i1} \beta_1 - X_{i2} \tilde{\beta}_1, \quad (1)$$

$$\text{logit}\{\Pr(Y_{i2k} \leq l | X_i, Z_{i2}, b_{i2})\} = \beta_{02l} - b_{i2}^T Z_{i2} - X_{i2} \beta_2 - X_{i1} \tilde{\beta}_2,$$

where $l = 1, \dots, L - 1$; intercepts β_{01l} and β_{02l} satisfy $\beta_{01l} < \dots < \beta_{01L-1}$, for $j = 1, 2$; b_{i1} and b_{i2} are subject-specific random effects with corresponding design matrixes Z_{i1} and Z_{i2} , which include 1 as the first column (for random intercepts). We intentionally use negative signs before regression coefficients β_1 , β_2 , $\tilde{\beta}_1$ and $\tilde{\beta}_2$ such that positive values of these parameters can be intuitively interpreted as that a larger value of covariate (i.e., X) leads to a higher probability that Y falls into higher categories. The model parameters in (1) can be interpreted using the actor-partner interdependence (API) framework, a conceptual framework widely used in psychology and the social sciences to explain intra-dyad interdependence (Cook and Kenny, 2005). Under the API framework, β_1 and β_2 represent the “actor” effect, which describes how a participant’s outcome is affected by his/her own covariates or characteristics, and $\tilde{\beta}_1$ and $\tilde{\beta}_2$ represent the “partner” effect, which describes

how a participant's outcome is affected by his/her partner's covariates or characteristics. The partner effect explains within-dyad interdependence induced by measured covariates. For example, in our study, $\tilde{\beta}_1$ describes how MBC patient's covariates affect her spouse's outcome. We capture the extra within-dyad interdependence that cannot be explained by the partner effect (e.g., caused by unmeasured covariates) through specifying a joint distribution of random effects $\mathbf{b}_{i1}$ and $\mathbf{b}_{i2}$, as expressed below.

There are two popular options for the ordinal responses such as a proportional odds model and a baseline-category *logit* model (Agresti, 2012). For example, when the monotone trend is not clear, the baseline-category *logit* model can be used by

$$\log \left(\frac{\Pr(Y_{i1k} = l | X_i, Z_{i1}, \mathbf{b}_{i1})}{\Pr(Y_{i1k} = L | X_i, Z_{i1}, \mathbf{b}_{i1})} \right) = \beta_{01l} - \mathbf{b}_{i1}^\top \mathbf{Z}_{i1} - X_{i1} \beta_{1l} - X_{i2} \tilde{\beta}_{1l}$$

$$\log \left(\frac{\Pr(Y_{i2k} = l | X_i, Z_{i2}, \mathbf{b}_{i2})}{\Pr(Y_{i2k} = L | X_i, Z_{i2}, \mathbf{b}_{i2})} \right) = \beta_{02l} - \mathbf{b}_{i2}^\top \mathbf{Z}_{i2} - X_{i2} \beta_{2l} - X_{i1} \tilde{\beta}_{2l}$$

where $l = 1, \dots, L - 1$. In the MBC study, the cumulative *logit* model is chosen over the baseline-category *logit* model due to ordinality of the response. The further details about this choice will be discussed in Section 5.

In longitudinal dyadic data, two types of correlations need to be taken into account: the within-subject correlation derived from the repeated measurements, and the within-dyad correlation between the two members of a dyad. This can be done through specifying a joint distribution of $\mathbf{b}_{i1}$ and $\mathbf{b}_{i2}$. We assume that $\mathbf{b}_{i1}$ and $\mathbf{b}_{i2}$ follow a multivariate normal distribution of the form,

$$\begin{pmatrix} \mathbf{b}_{i1} \\ \mathbf{b}_{i2} \end{pmatrix} \sim MVN \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \mathbf{\Lambda} \right), \quad (2)$$

where $\mathbf{\Lambda}$ is a covariance matrix. In the case that $\mathbf{b}_{ij}$ contains only two elements, namely a random intercept b_{ij0} and a random slope b_{ij1} , a general form of covariance $\mathbf{\Lambda}$ can be expressed as

$$\mathbf{\Lambda} = \begin{pmatrix} \nu_{10}^2 & \nu_{10}\nu_{11}\rho_1 & \nu_{10}\nu_{20}\rho_3 & \nu_{10}\nu_{21}\rho_4 \\ \cdot & \nu_{11}^2 & \nu_{11}\nu_{20}\rho_5 & \nu_{11}\nu_{21}\rho_6 \\ \cdot & \cdot & \nu_{20}^2 & \nu_{20}\nu_{21}\rho_2 \\ \cdot & \cdot & \cdot & \nu_{21}^2 \end{pmatrix}, \quad (3)$$

where ν_{j0}^2 and ν_{j1}^2 are the variances of the random intercept and slope, respectively, for member j , where $j = 1, 2$. For example, with patient ($j = 1$) and spouse ($j = 2$) in the MBC study, there are six correlations: ρ_1 (or ρ_2) captures the correlation between random intercepts and random slopes from the patient (or spouse) in a dyad in the MBC study; ρ_3 (or ρ_4) accounts for the correlation between the patient's random intercepts and the spouse's random intercepts (or slopes), and lastly ρ_5 (or ρ_6) explains the correlation between the patient's random slopes and the spouse's random intercepts (or slopes). The latter four correlations account for within-dyad interdependence.

The multivariate normal assumption for $(\mathbf{b}_{i1}, \mathbf{b}_{i2})$ can be relaxed by assuming that $(\mathbf{b}_{i1}, \mathbf{b}_{i2})$ follow a multivariate t distribution, which provides more robust inference (Lange, et al., 1989),

$$\begin{pmatrix} \mathbf{b}_{i1} \\ \mathbf{b}_{i2} \end{pmatrix} \sim t_k \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \mathbf{\Psi} \right), \quad (4)$$

where $\mathbf{\Psi}$ is the covariance matrix and κ is the degree of freedom. In addition, we can also relax the homoscedastic assumption in (3) by letting the covariance matrix $\mathbf{\Lambda}$ depend on covariates $\mathbf{X}_j$. We demonstrate these alternative approaches in the analysis of MBC data in Section 5.

In the absence of missing data, the complete-data likelihood conditional on the random effects is given by

$$f(\mathbf{Y}|\boldsymbol{\theta}_1, \mathbf{X}, \mathbf{b}) = \prod_{i=1}^n \prod_{j=1}^2 \prod_{k=1}^K \prod_{l=1}^L f(y_{ijk} | \mathbf{b}_{ij}, \mathbf{X}_{ijk}, \boldsymbol{\theta}_1) = \prod_{i=1}^n \prod_{j=1}^2 \prod_{k=1}^K \prod_{l=1}^L \left[\frac{\exp(\beta_{0jl} - \mathbf{b}_{ij}^\top \mathbf{Z}_{ij} - \mathbf{X}_{ij} \boldsymbol{\beta}_j - \tilde{\mathbf{X}}_{ij} \tilde{\boldsymbol{\beta}}_j)}{1 + \exp(\beta_{0jl} - \mathbf{b}_{ij}^\top \mathbf{Z}_{ij} - \mathbf{X}_{ij} \boldsymbol{\beta}_j - \tilde{\mathbf{X}}_{ij} \tilde{\boldsymbol{\beta}}_j)} - \frac{\exp(\beta_{0jl-1} - \mathbf{b}_{ij}^\top \mathbf{Z}_{ij} - \mathbf{X}_{ij} \boldsymbol{\beta}_j - \tilde{\mathbf{X}}_{ij} \tilde{\boldsymbol{\beta}}_j)}{1 + \exp(\beta_{0jl-1} - \mathbf{b}_{ij}^\top \mathbf{Z}_{ij} - \mathbf{X}_{ij} \boldsymbol{\beta}_j - \tilde{\mathbf{X}}_{ij} \tilde{\boldsymbol{\beta}}_j)} \right]^{I(y_{ijk}=l)},$$

where $\boldsymbol{\theta}_1 = (\beta_{01}, \beta_{02}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2)$ is the vector of unknown parameters.

2.3 Transition model for intermittent data missingness and dropouts

To model the missing data process, we should account for the fact that the missing data processes for the two members of a dyad typically are correlated. If one member of a dyad misses a scheduled visit or drops out of the study, the other member tends to do so as well. On the other hand, we should also allow for the flexibility such that two members of a dyad can drop out at different times. In our motivating data, we observe that in some dyads, two members dropped out at different time points and had different patterns of intermittent data missingness. We should also account for the possibility that dropping out of a study and having intermittent missing data may be caused by different missing data mechanisms and that the two members of a dyad also may have different missing data mechanisms.

To this end, define $\pi_{ijk}^{(\mathcal{O})} = \Pr(D_{ijk} = \mathcal{O} | D_{ij(k-1)} \neq \mathcal{D}, X_{ij}, \mathbf{b}_{ij})$ as the transition probability from status $\mathcal{O}$ (observed) or $\mathcal{J}$ (intermittently missing data) at the $(k-1)$ th measurement time to $\mathcal{O}$ at the k th measurement time, for a participant with covariates X_{ij} and random effect $\mathbf{b}_{ij}$. Similarly, define $\pi_{ijk}^{(\mathcal{J})} = \Pr(D_{ijk} = \mathcal{J} | D_{ij(k-1)} \neq \mathcal{D}, X_{ij}, \mathbf{b}_{ij})$ as the transition probability from $\mathcal{O}$ or $\mathcal{J}$ to $\mathcal{J}$, and define $\pi_{ijk}^{(\mathcal{D})} = \Pr(D_{ijk} = \mathcal{D} | D_{ij(k-1)} = \mathcal{O}, X_{ij}, \mathbf{b}_{ij})$ as the transition probability from $\mathcal{O}$ to $\mathcal{D}$ (dropout). We propose the following proportional odds transition model to describe the missing data process,

$$\pi_{ijk}^{(\mathcal{O})} = \frac{1}{1 + \sum_{m \in \{\mathcal{J}, \mathcal{D}\}} \exp(d_i + X_{ij}^T \boldsymbol{\eta}_j^{(m)} + \mathbf{b}_{ij}^T \boldsymbol{\xi}_j^{(m)})}, \quad (5)$$

$$\pi_{ijk}^{(\mathcal{J})} = \frac{\exp(d_i + X_{ij}^T \boldsymbol{\eta}_j^{(\mathcal{J})} + \mathbf{b}_{ij}^T \boldsymbol{\xi}_j^{(\mathcal{J})})}{1 + \sum_{m \in \{\mathcal{J}, \mathcal{D}\}} \exp(d_i + X_{ij}^T \boldsymbol{\eta}_j^{(m)} + \mathbf{b}_{ij}^T \boldsymbol{\xi}_j^{(m)})},$$

$$\pi_{ijk}^{(\mathcal{D})} = \frac{\exp(d_i + X_{ij}^T \boldsymbol{\eta}_j^{(\mathcal{D})} + \mathbf{b}_{ij}^T \boldsymbol{\xi}_j^{(\mathcal{D})})}{1 + \sum_{m \in \{\mathcal{J}, \mathcal{D}\}} \exp(d_i + X_{ij}^T \boldsymbol{\eta}_j^{(m)} + \mathbf{b}_{ij}^T \boldsymbol{\xi}_j^{(m)})},$$

where d_i is the dyad-specific random effect accounting for the correlation between the two members' missing data processes, such that if one member drops out or misses a measurement, the other member tends to also drop out or miss the measurement. We assume $d_i \sim N(d, \sigma_d^2)$ where d is the expected value of the intercept estimator. The parameter vector $\boldsymbol{\eta}_j^{(m)}$ represents the covariate effects of $\mathbf{X}$ on the transition probabilities. The parameter vector $\boldsymbol{\xi}_j^{(m)}$ describes how the subject-specific random effects $\mathbf{b}_{ij}$ affect the missing data process. By sharing $\mathbf{b}_{ij}$ between the outcome model and the missing data model, we link the two processes together and thus account for (potentially) non-ignorable missing data. The value of $\boldsymbol{\xi}_j^{(m)}$, $m = \mathcal{J}, \mathcal{D}$, controls the missing data mechanisms of intermittent data missingness and dropout, respectively. Specifically, when $\boldsymbol{\xi}_j^{(I)} = 0$ (or $\neq 0$), the intermittent data missingness is ignorable (or non-ignorable); when $\boldsymbol{\xi}_j^{(D)} = 0$ (or $\neq 0$), dropout is ignorable (or non-ignorable). Note that in our case, modeling the above three transition probabilities is adequate. This is because by definition, $\mathcal{D}$ (dropout) is an absorbing state, and thus the transition probabilities from $\mathcal{D}$ to other states are 0. In addition, we require that $\Pr(D_{ijk} = \mathcal{D} | D_{ij(k-1)} = \mathcal{J}) = 0$ because the intermittent data missingness, by definition, cannot be immediately followed by a dropout. Here we assume that the outcome at the first visit (i.e., baseline) is observed for all participants. To obtain reliable estimates of the transition model, a reasonable number of observations should be available for each transition (or missing data) pattern. In the MBC data, there are 27 and 25 observations with

intermittent missing pattern ($\mathcal{J}$) for patients and spouses, and 97 and 98 observations with drop-out pattern ($\mathcal{D}$) for patients and spouses, respectively (see Table 1). Our data analysis later shows that these data provide sufficient information to obtain reliable parameter estimates with reasonably wide credible intervals (see Table 3). In practice, with the scarcity of data in a specific missing pattern, the estimation of the missing mechanism parameters can be unstable or unidentifiable. In such a situation, one can collapse two separate missing mechanisms ($\mathcal{D}$) and ($\mathcal{J}$) into one missing data model that requires a less number of parameters to estimate. In doing so, one may consider an indicator variable that differentiates the two missing patterns.

Conditional on random effects $\mathbf{b}_{ij}$, the likelihood of $\mathbf{D}_{ij}$ contributed by participant i of dyad j is

$$f(\mathbf{D}_{ij} | \mathbf{X}_i, \mathbf{b}_{ij}, \boldsymbol{\theta}_2) = \prod_{k=2}^K \pi_{ijk}^{(\mathcal{O})} I(D_{ijk} = \mathcal{O}) \pi_{ijk}^{(\mathcal{J})} I(D_{ijk} = \mathcal{J}) \pi_{ijk}^{(\mathcal{D})} I(D_{ijk} = \mathcal{D}),$$

where $\boldsymbol{\theta}_2 = \{\mathbf{d}, \boldsymbol{\eta}_j^{(m)}, \boldsymbol{\xi}_j^{(m)}\}$ denotes the vector of the parameters for the transition process, and after participant i of the j th dyad drops out, the rest of the D_{ijk} 's are undefined. Let $\mathbf{Y}_{ij}^{\text{obs}}$ represent the vector of the observed outcomes and $\mathbf{Y}_{ij}^{\text{mis}}$ represent the vector of the unobserved outcomes. The joint likelihood of the outcome and dropout processes is given by

$$\begin{aligned} f(\mathbf{D}, \mathbf{Y}^{\text{obs}} | \mathbf{X}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= \int \prod_{i=1}^n \left\{ \prod_{j=1}^2 f(\mathbf{Y}_{ij}^{\text{obs}}, \mathbf{Y}_{ij}^{\text{mis}} | \mathbf{X}_{ijk}, \mathbf{b}_{ij}, \boldsymbol{\theta}_1) f(\mathbf{D}_{ij} | \mathbf{X}_i, \mathbf{b}_{ij}, \boldsymbol{\theta}_2) d\mathbf{Y}_{ij}^{\text{mis}} \right\} f(\mathbf{b}_{ij}) d\mathbf{b}_{ij} \\ &= \int \prod_{i=1}^n \left\{ \prod_{j=1}^2 \prod_{k \in \{k | D_{ijk} = \mathcal{O}\}} f(\mathbf{Y}_{ijk} | \mathbf{X}_{ijk}, \mathbf{b}_{ij}, \boldsymbol{\theta}_1) f(\mathbf{D}_{ij} | \mathbf{X}_i, \mathbf{b}_{ij}, \boldsymbol{\theta}_2) \right\} f(\mathbf{b}_{ij}) d\mathbf{b}_{ij}. \end{aligned}$$

3 Prior and posterior inference

Under the Bayesian paradigm, we assign independent vague priors to model parameters as follows,

$$\beta_{01}, \beta_{02} \sim N(0, q^{-1}) I(\beta_{0j1} < \dots < \beta_{0jL-1}), \quad j = 1, 2,$$

$$\beta_1, \tilde{\beta}_1, \beta_2, \tilde{\beta}_2 \stackrel{iid}{\sim} N(0, q^{-1}),$$

$$d_i, \eta_1^{(\mathcal{J})}, \eta_2^{(\mathcal{J})}, \eta_1^{(\mathcal{D})}, \eta_2^{(\mathcal{D})}, \xi_1^{(\mathcal{J})}, \xi_2^{(\mathcal{J})}, \xi_1^{(\mathcal{D})}, \xi_2^{(\mathcal{D})} \stackrel{iid}{\sim} N(0, q^{-1}),$$

$$\Lambda^{-1} \sim WI_{\nu_0}(cI),$$

$$\sigma_d^2 \sim IG(\sigma_a, \sigma_b),$$

where $WI_{\nu_0}(cI)$ represents the Wishart distribution with ν_0 degrees of freedom and a scale matrix cI . We assigned small constant values, 10^{-6} , for ν_0 and c . $IG(\sigma_a, \sigma_b)$ denotes an inverse gamma distribution with shape parameter σ_a and scale parameter σ_b . In order to minimize the influence of the priors on the estimation, we set the precision $q = 10^{-6}$ for normal priors $N(0, q^{-1})$ and $\sigma_a = \sigma_b = 10^{-6}$. For β_{01} and β_{02} , the indicator function $I(\cdot)$ is used to ensure an appropriate ordering for intercepts. Let $p(\theta_1, \theta_2)$ denote the above prior distributions, the posterior is given by

$$p(\theta_1, \theta_2 | D, Y^{\text{obs}}, X) = p(\theta_1, \theta_2) f(D, Y^{\text{obs}} | X, \theta_1, \theta_2).$$

We sample this posterior using the Gibbs sampler with the Metropolis Hastings algorithm. Details of the Gibbs sampler are provided in Supplementary Materials.

4 Simulation

We conducted a simulation study to compare the performance of the proposed DSPM with that of a dyadic generalized linear mixed model (DGMM) given by equations (1) and (2). The DGMM is a simplified version of the DSPM by ignoring the missing data process. For the purpose of comparison, we also implemented a before-deletion version of the DGMM (denoted as DGMM-BD), i.e., fitting the DGMM based on the complete data before deleting the missing data. Although the DGMM-BD is not available in practice, it provides a benchmark for evaluating the DSPM and DGMM in a simulation study. Lastly, to appreciate the importance of accounting for the within-dyad interdependence, we also fit a standard generalized linear mixed model before deleting the missing data (denoted as GMM-BD), which is the same as the DGMM-BD except that GMM-BD assumes that b_{j1} and b_{j2} are independent. The comparison between the GMM-BD and DGMM-BD informs us of the consequence of ignoring the within-dyad interdependence.

We assumed $n = 200$ dyadic pairs and that the ordinal outcome Y_{ijk} ($= 1, 2, 3$, or 4) was repeatedly measured over four time points (i.e., $Z_{ij} = 1, 2, 3$, or 4). We considered a single covariate X_1 for the first member and X_2 for the second member, both of which were simulated from $N(1, 1)$. We simulated the outcome Y_{ijk} based on the outcome process model (1), with $(\beta_1, \tilde{\beta}_1) = (\beta_2, \tilde{\beta}_2) = (2, -1)$ and $(\beta_{011}, \beta_{012}, \beta_{013}) = (\beta_{021}, \beta_{022}, \beta_{023}) = (-2, 0, 2)$, such that four levels of Y_{ijk} had approximately the same number of observations. We simulated random effects, which consisted of a random intercept and a random slope, from a multivariate normal distribution with a mean vector $(0, 1, 0, 1)$, a correlation vector $(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6) = (0.2, 0.2, 0.2, 0.2, 0.2, 0.2)$, and variances $\mathbf{v} = (2, 2, 2, 2)$ given model (3).

We considered eight scenarios with different types of missing data mechanism, among which the first six are defined based on transition model (5).

(Sc. 1) Both dropout and intermittent missing data are ignorable (i.e., missing completely at random) by setting $\eta^{(\mathcal{D})} = \zeta^{(\mathcal{D})} = \eta^{(\mathcal{I})} = \zeta^{(\mathcal{I})} = 0$.

(Sc. 2) Dropout is non-ignorable, but intermittent missing data are ignorable by setting $(\zeta_1^{(\mathcal{D})}, \zeta_2^{(\mathcal{D})}) = (0, 1)$ and $(\zeta_1^{(\mathcal{I})}, \zeta_2^{(\mathcal{I})}) = (0, 0)$.

(Sc. 3) Both dropout and intermittent missing data are non-ignorable by setting $(\zeta_1^{(\mathcal{I})}, \zeta_2^{(\mathcal{I})}) = (0, 1)$ and $(\zeta_1^{(\mathcal{D})}, \zeta_2^{(\mathcal{D})}) = (0, 1)$.

(Sc. 4) Dropout is non-ignorable, but intermittent missing data are ignorable by setting $(\zeta_1^{(\mathcal{D})}, \zeta_2^{(\mathcal{D})}) = (1, 0)$ and $(\zeta_1^{(\mathcal{I})}, \zeta_2^{(\mathcal{I})}) = (0, 0)$.

(Sc. 5) Both dropout and intermittent missing data are non-ignorable by setting $(\zeta_1^{(\mathcal{I})}, \zeta_2^{(\mathcal{I})}) = (1, 0)$ and $(\zeta_1^{(\mathcal{D})}, \zeta_2^{(\mathcal{D})}) = (1, 0)$.

(Sc. 6) Both dropout and intermittent missing data are non-ignorable, but the transition model is mis-specified, with X_{ij} being replaced by a quadratic term X_{ij}^2 in the model (5). This represents a moderate model misspecification because the correlation between X_{ij1} and X_{ij1}^2 is about 0.82. We set $(\zeta_1^{(\mathcal{I})}, \zeta_2^{(\mathcal{I})}) = (0, 1)$ and $(\zeta_1^{(\mathcal{D})}, \zeta_2^{(\mathcal{D})}) = (0, 1)$.

(Sc. 7) Both dropout and intermittent missing data are non-ignorable, but the probability of missingness directly depends on the unobserved outcome Y_{ijk} in the form of $\text{logit}(\Pr(D_{ijk} = \mathcal{D} \text{ or } \mathcal{I})) = 0.5 + 0.5Y_{ijk}$. In this case, the transition model assumption is severely violated.

(Sc. 8) Both dropout and intermittent missing data are non-ignorable with the probability of missingness depending on both the unobserved outcome Y_{ijk} and the random intercept b_{ij1} in the form of $\text{logit}(\Pr(D_{ijk} = \mathcal{D} \text{ or } \mathcal{I})) = 0.5 + 0.5Y_{ijk} + 0.5b_{ij1}$. In this case, the transition model assumption is also violated.

The dyad-specific random effect d_i was simulated from $\mathcal{N}(1, 1)$ herein. We adjust transition model parameters such that approximately 25%, 30%, and 35% missing data occurred at the second, third, and the last visits, respectively. For example, we set $\eta_1^{(\mathcal{D})} = 1$ in scenarios 2 to 6, and $\eta_1^{(\mathcal{I})} = 0$ in scenario 2 and 4 and $\eta_1^{(\mathcal{I})} = 1$ in scenarios 3, 5, and 6. As a result, there are sufficient numbers of observations (at least 30) for each missing data or transition pattern for each dyadic member. We conducted 1,000 simulations under each of the six scenarios of missing data.

The proposed DSPM makes a variety of parametric assumptions, e.g., multivariate normal distributions for random effects with a homoscedastic covariance matrix, and linear mean structures for both outcome and transition models. We further investigated the impact of

violations of such model assumptions through sensitivity analyses in Supplementary materials.

Table 2 reports the simulation results from scenarios 1 through 8, including the relative bias (RB, bias/true value), mean squared error (MSE), and coverage probability (CP) of the parameter estimates across the 1,000 replicates. As expected, the DGMM-BD had the least bias, with coverage probabilities close to the 95% nominal level. Compared to this reference standard, the GMM-BD, which is also based on the complete data but ignores the dyadic interdependence, led to biased estimates with poor coverage probabilities of approximately 20% for β_1 and β_2 . The low coverage probabilities indicate that under repeated sampling, the 95% credible interval has a low chance to cover the true values of the parameter. These results highlight the importance of modeling the dyadic interdependence.

The proposed DSPM performed well under different missing data mechanisms. The resulting estimates had minimal bias and sound coverage probabilities that were close to nominal 95% under scenarios 1, 2, 3, 4, and 5. These estimates were less efficient than those based on the DGMM-BD, with larger MSEs because of the missing data. In contrast, the DGMM performed reasonably well only when the dropout data and intermittent missing data are ignorable (i.e., scenario 1). When the dropout data are non-ignorable but the intermittent missing data are ignorable (i.e., scenario 2), the DGMM produced larger biases than the DSPM, and the coverage probability for some estimates (e.g., estimate of β_1 and β_2) were lower than 90%. Similar results were observed under scenario 3 when both the dropout and intermittent missing data were non-ignorable. When the random intercepts affect the missing mechanisms (i.e., scenario 4 and 5), the DGMM produced more biases in β_1 and β_2 compared to the DGMM under scenario 2 and 3 while the DSPM performs well equally. Scenario 6 was constructed to evaluate the sensitivity of the methods when the missing data mechanism is moderately misspecified. We can see that the estimates based on the DSPM had slightly larger biases and lower coverage than those under scenarios 1 to 5, but they are still substantially better than the estimates of the DGMM. In scenarios 7 and 8, the missing mechanism directly depends on unobserved outcomes, and the assumed (random-effects based) missing data mechanism is strongly violated. The DSPM still outperformed the DGMM with smaller biases and MSEs.

5 Application

5.1 Model fitting

We applied the proposed DSPM to analyze the metastatic breast cancer (MBC) data. The outcome of interest is the severity of depression categorized as three levels: 1=no depression (63%), 2=mild depression (22%), and 3=major depression (15%). After consulting with our collaborator, four covariates that were expected to be potentially correlated with the outcome were included in the outcome model: the patient's chronic pain as measured by the Multidimensional Pain Inventory (MPI, seven levels with 0 being for the least pain and 6 for the most pain imaginable), spouse's chronic pain as measured by the MPI, the measurement time (i.e., baseline, 3 months and 6 months, coded as 0, 1, 2). We determined to use fixed intercepts because the model with random intercepts led to a convergence problem whilst random slopes for the time effect did not. To this end, a simpler version of covariance 2×2

structure with non-null $\rho_6, \nu_{11}^2, \nu_{21}^2$ was employed rather than model 3. We preliminary considered two candidate outcome models: the cumulative proportional odds model and baseline-category model. The deviance information criterion (DIC, Spiegelhalter et al., 2002), a goodness-of-fit statistic, suggests in favor of the proportional odds model (DIC = 592) over the baseline-category *logit* model (DIC = 602). Therefore, we employed the proportional odds model to analyze the data. In the transition model for the missing data process, we included the subject's chronic pain as a sole covariate because it was expected that the probability that a subject missed a visit or dropped out of the study should largely depend on his/her own characteristics (i.e., covariates and outcome trajectory), not his/her spouse's covariates. This is also because the number of observations in each transition pattern is limited and the generous transition model had led to unstable estimates.

Considering the complexity of the model, we ran 50,000 MCMC iterations to fit the model and discarded the first 20,000 as burn-in. And after thinning every 10 posterior draws, we made inferences based on the remaining 3,000 posterior samples. We monitored the convergence of the MCMC algorithm based on the trace plot and the convergence diagnostic proposed by Gelman and Rubin (1992). The trace plot shows that the MCMC chain was well mixed, and convergence diagnostic of Gelman and Rubin (1992) was generally close to 1 (ranging from 1.0 to 1.03), suggesting convergence of the MCMC algorithm.

Table 3 reports the parameter estimates and 95% highest posterior density (HPD) credible intervals (CIs) for the missing data model. The results suggest that for patients, dropout data are likely to be non-ignorable as the 95% credible interval of $\zeta_{11}^{(\mathcal{D})}$ exclude 0. For both patient and spouse, the intermittent missing data and dropout data may involve different missing data mechanisms. The significantly negative estimates of $\eta_{11}^{(\mathcal{F})}$ (95% CI of (-1.40, -0.64)) and $\eta_{21}^{(\mathcal{F})}$ (95% CI of (-1.42, -0.72)) for the patients and spouse in the intermittent missingness process suggest that when either patient or spouse has a higher MPI, both members of the dyad tend to adhere to the study without any delinquency. This trend is relatively weak for the dropout missingness with 95% CI barely excluding the null association.

Table 4 shows the parameter estimates in the outcome process. For comparison, we also display the results when fitting the data using the DGMM, which assumes that the missing data are ignorable or missing at random. Both the DSPM and DGMM suggest significant actor and partner effects for the MPI on depression for patients, i.e., chronic pain experienced by either the patient or the spouse is associated with depression in the person who is directly experiencing the pain. A more severe level of depression is more likely for the patient or spouse who has a higher MPI. Both the DSPM and DGMM capture this relationship quite similarly. For example, for patients, the difference in estimates of the actor and partner effects under the DSPM (i.e., 0.41-0.15=0.26) is very close to that under the DGMM (i.e., 0.41-0.16=0.25). The estimate of the effect of the time under the DSPM is rather different from that obtained under the DGMM. For example, for the spouse, the estimate under the DSPM is -0.16 with 95% CI=(-0.47, 0.11) including the null; whereas the estimate under the DGMM is -0.46 with 95% CI = (-0.79, -0.13) not including the null.

Such differences may be due to the fact that the DGMM ignores the potentially non-ignorable missing data.

As the estimates of the outcome process are generally similar between the DGMM and DSPM, one may question the value of the DSPM over the DGMM and wonder what does the DSPM offer that the DGMM does not, for this specific application. The MBC data suffered from a substantial amount (approximately 40%) of missing data, and the investigators had worried about the estimation bias due to potentially non-ignorable missing data. If we only fit the DGMM, we would have no clue whether it is reasonable to ignore the missing data and whether the analysis results are biased due to ignoring the missing data. By fitting the DSPM and comparing the results between the DSPM and DGMM, we can answer these important practical questions. In this specific case, it turns out that ignoring missing data does not substantially affect the results. This provides us additional confidence and reassurance on the analysis results. Also, the DSPM here is still valuable as it provides extensive sensitivity analyses that confirm the robustness of the DSPM estimates under the moderate misspecification of the model assumptions (See Supplementary Material). In addition, the DSPM sheds lights on the missing data mechanism of the dropouts and intermittent missing data, rendering the investigators better understanding of the missing data process for the study.

Table 5 displays the estimates and 95% CI of the covariance components for the random effects model under the DSPM. The results suggest a significant positive correlation between the random slopes from patients and spouses, with $\hat{\rho}_6 = 0.12$ (95% CI=(0.02, 0.22)), which demonstrates the within-dyad interdependence. On the other hand, within-dyad interdependence is not significant under the DGMM with $\hat{\rho}_6 = 0.05$ (95% CI=(-0.07, 0.16). The estimate (95% HPD credible intervals) for random effect variance σ^2 for d_i is 11.03(6.51, 15.74).

5.2 Model assessment

We assessed the goodness of fit of the proposed model using the posterior predictive checking method, which is based on the idea that if the model fits the data, the observed data should be consistent with the replicated data generated from the proposed model. We adopted the following Bayesian omnibus residual statistics to summarize the discrepancy between the values predicted under the fitted model and the data that were actually observed (Gelman et al., 2013),

$$T(\mathbf{y}, \boldsymbol{\theta}) = \sum_i^n \sum_j^2 \sum_{k \in \{k | D_{ijk} = \emptyset\}} \frac{\{y_{ijk} - E(y_{ijk} | \boldsymbol{\theta})\}^2}{\text{var}(y_{ijk} | \boldsymbol{\theta})}.$$

We drew S replicates of the data $\mathbf{y}$, denoted as $\{\mathbf{y}^{rep,s}, s = 1, \dots, S\}$, from the posterior predictive distribution $p(\mathbf{y}^{rep} | \mathbf{y})$ through simulations, and then computed the Bayesian p-value as the proportion of $T(\mathbf{y}^{rep,l}, \boldsymbol{\theta})$ that is equal to or larger than $T(\mathbf{y}, \boldsymbol{\theta})$ out of S replicates. The estimated p-value is 0.782, which suggests that the model fits the data well.

The residual Figure 1 in Supplementary Material also shows no major problems appearing on this fit.

Our model makes a proportional odds assumption on the covariate effect, that is, the covariate effect is the same across different outcome categories. We examined that assumption by fitting the data with the following model,

$$\text{logit}\{\Pr(Y_{i1k} \leq l | X_i, Z_{i1}, b_{i1})\} = \beta_{01l} - b_{i1}^T Z_{ij} - X_{i1} \beta_{1l} - X_{i2} \tilde{\beta}_{1l}$$

$$\text{logit}\{\Pr(Y_{i2k} \leq l | X_i, Z_{i1}, b_{i1})\} = \beta_{02l} - b_{i2}^T Z_{ij} - X_{i2} \beta_{2l} - X_{i1} \tilde{\beta}_{2l}, \quad l = 1, 2.$$

Compared to the proportional odds model (1), the above model is more flexible because it allows the covariate effect to vary across outcome categories as β_{jl} and $\tilde{\beta}_{jl}$ are now indexed by outcome category l . Noting that if $\beta_{j1} = \beta_{j2}$ and $\tilde{\beta}_{j1} = \tilde{\beta}_{j2}$ (or equivalently $\beta_{j1} - \beta_{j2} = 0$ and $\tilde{\beta}_{j1} - \tilde{\beta}_{j2} = 0$) for $j = 1, 2$, the above model becomes the proportional odds model. We calculated the 95% highest posterior density (HPD) intervals for $\beta_{j1} - \beta_{j2}$ and $\tilde{\beta}_{j1} - \tilde{\beta}_{j2}$, for $j = 1, 2$. As shown in Supplementary Table 1, all the HPD intervals contained 0, suggesting that the covariate effect does not significantly vary across the outcome categories. That is, there is no significant evidence against the proportional odds assumption.

Another important assumption underlying the proposed DSPM is that conditional on random effect b_{ij} , the missing data process is independent of Y^{mis} , the missing values of Y . We checked this conditional independence assumption by refitting the missing data model (5) and adding the last observed value of Y (say Y^{last}) as an extra covariate, and then examining whether Y^{last} is significantly associated with the missing data probabilities after controlling for random effects b_{ij} in the model. Ideally, we should include Y^{mis} as the covariate in order to examine whether the missing data process is conditionally independent of the missing values of Y . However, that approach is computationally challenging because Y^{mis} is not observed. Thus, we took an approximate approach by using Y^{last} as the surrogate of Y^{mis} , motivated by the fact that in practice the last observed value of Y is often highly correlated with the subsequent (missing) measurement. The results show that Y^{last} is not significantly associated with the transition process D . The 95% CIs are $(-0.42, 0.23)$, $(-0.82, 0.43)$, $(-0.51, 0.32)$ and $(-0.42, 0.23)$ for the two members under the two types of missing data (i.e., $\mathcal{D}$ or $\mathcal{F}$). This evidence supports empirically that the conditional independence assumption may be reasonable for the MBC data.

The proposed DSPM assumes that random effects (b_{j1}, b_{j2}) follow a multivariate normal distribution with a homoscedastic covariance matrix. To confirm these parametric assumptions are reasonable, we fitted two alternative models that 1) assumed a multivariate t distribution with degrees of freedom 1 or 5 for (b_{j1}, b_{j2}) ; and 2) assumed that the covariance matrix (3) depends on actor's MPI values. The estimates of the parameters of main interest for 1) and 2) are reported in Supplementary Table 4 and 5, respectively. We found that the

estimates were generally similar between the DSPM and these more flexible models, suggesting the parametric assumptions are reasonable for this application. We also investigated the impact of prior specification of β , η , ζ on the inference by using different values of hyper-parameters (i.e., the value of q and σ_a and σ_b). The results (see Supplementary Table 6) are almost identical to these reported in Table 4, suggesting that our prior distributions are sufficiently non-informative.

5.3 Sensitivity analysis for the proposed missing mechanism

A fundamental inherent issue when modeling non-ignorable missing data is that the assumption of the missing data mechanism is not testable based on the observed data and thus the identification of some model parameters is largely driven by the model assumption. A common approach to address this issue is sensitivity analysis, which examines the sensitivity of the estimates of primary interest with respect to the values of the parameters that control the missing data mechanism (Rotnitzky, Robins and Scharfstein, 1998; Daniels and Hogan, 2000, 2008). In our case, $\zeta^{(\mathcal{D})}$ and $\zeta^{(\mathcal{F})}$ control the missing data mechanism. We fixed $\zeta^{(\mathcal{D})}$ and $\zeta^{(\mathcal{F})}$ as a series of fixed values (rather than estimating them from the data) and investigated how the estimates of the other parameters (based on the data) vary with respect to the values of $\zeta^{(\mathcal{D})}$ and $\zeta^{(\mathcal{F})}$.

Figure 1 depicts the changes in the estimates for the outcome model when $\zeta_{11}^{(\mathcal{D})} = \zeta_{21}^{(\mathcal{F})} = \zeta_{11}^{(\mathcal{D})} = \zeta_{21}^{(\mathcal{F})}$ change from -2 to 2 . This range will approximately make the effect size of $b_{ij}\zeta_j^{(m)}$ in (5) vary between -4 and 4 with 95% given estimates of ν in Table 5.

In general, the estimates for the outcome model are seemingly stable when $\zeta^{(\mathcal{D})}$ and $\zeta^{(\mathcal{F})}$ change, which suggests a limited impact on our estimates and the parameter estimation under our model is reasonably robust when the missing data model parameters are misspecified.

6 Conclusion

In this article, we propose the DSPM to analyze longitudinal dyadic data with non-ignorable dropouts and intermittent missing data. By allowing the outcome process and dropout process to share common dyad-specific random effects, the DSPM accounts for the informative missing data. Dyadic interdependence is captured through additional random effects in the outcome model and missing data model, as well as the actor-partner effect parameters. Simulation studies show that the proposed model yields minimally biased estimates in the presence of informative missing data, and these estimates are robust to some degree of model misspecification on the missing data mechanism. It is worth noting that, as all methods dealing with non-ignorable missing data, our approach does not solve the intrinsic non-identification problem of non-ignorable missing data, and its validity strongly depends on how plausible untestable missing data mechanism is the assumed. When using the proposed method to analyze longitudinal dyadic data with missing data, it is imperative to evaluate the plausibility of the underlying model assumptions by consulting with subject matter experts, and conduct extensive sensitivity analyses to assess how the analysis results change when the missing data mechanism is misspecified.

We agree that in our MBC case study, the estimates of the outcome process are similar between the DGMM and DSPM. This may be because that the missing data mechanism is nearly ignorable. Our simulation study shows that if the missing data mechanism is strongly non-ignorable, the estimates between the DGMM and DSPM can be rather different. One fundamental difficulty of analyzing potentially non-ignorable missing data is that the missing data mechanism is not testable. That is, based on the observed data, we cannot tell whether the missing data are ignorable or non-ignorable. Because of that, the modern literature on missing data analysis (Little et al., 2012) puts great emphasis on sensitivity analysis.

The proposed approach relies on some parametric assumptions. It is critical to assess the validity of these assumptions in applications. Alternatively, we can consider a more flexible class of shared-parameter models, such as semi-parametric approach that assumes a parametric model for the missing data mechanism and a nonparametric model for the outcome process. Smoothing splines can be used to specify the relationship between the outcome and the covariates without imposing a parametric structure on the mean. For the random effects, a multivariate t distribution or skewed multivariate normal distribution can be used to improve robustness, and various Bayesian non-parametric models (Hjort, et al., 2010) are available to further relax the distributional assumptions on the random effects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Agresti, A. Categorical data analysis. 3rd. Wiley; New York: 2012.
- Ahn J, Liu S, Wang W, Yuan Y. 2013; Bayesian latent class mixed effect hybrid models for dyadic longitudinal data with non ignorable dropouts. *Biometrics*. 69:914–924. [PubMed: 24328715]
- Badr H, Carmack CL, Kashy DA, Cristofanilli M, Revenson TA. 2010; Dyadic coping in metastatic breast cancer. *Health Psychology*. 29:169–180. [PubMed: 20230090]
- Carmack Taylor CL, Badr H, Lee L, Pisters K, Fossella F, Gritz ER, Schover L. 2008; Lung Cancer patients and their spouses: Psychological and relationship functioning within 1 month of treatment initiation. *Annals of Behavioral Medicine*. 36:129–140. [PubMed: 18797978]
- Carrell SE, Fullerton RL, West JE. 2009; Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*. 27:439–464.
- Cook WL, Kenny DA. 2005; The actor-partner interdependence model: a model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*. 29:101–109.
- Cook J, Tyson R, White J, Rushe, Gottaman JM, Murray JD. 1995; The mathematics of marital conflict: qualitative dynamic mathematical modeling of marital interaction. *Journal of Family Psychology*. 9:110–130.
- Daniels M, Hogan J. 2000; Reparameterizing the pattern mixture model for sensitivity analyses under informative dropout. *Biometrics*. 56:1241–1248. [PubMed: 11129486]
- Daniels, M, Hogan, J. Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis. Monographs on Statistics and Applied Probability 109 Chapman & Hall/CRC; Boca Raton, FL: 2008.
- Diggle P, Kenward MG. 1994; Informative drop-out in longitudinal data analysis. *Applied Statistics*. 43:49–73.

- Fang CY, Manne S, Pape SJ. 2001; Functional impairment, marital quality, and patient psychological distress as predictors of psychological distress among cancer patients' spouses. *Health Psychology*. 20:452–457. [PubMed: 11714188]
- Follman D, Wu MC. 1995; An approximate generalized linear model with random effects for informative missing data. *Biometrics*. 51:151–168. [PubMed: 7766771]
- Gardner M, Steinberg L. 2005; Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study. *Developmental Psychology*. 41:625–635. [PubMed: 16060809]
- Gelman, A, Carlin, J, Stern, H, Dunson, D, Vehtari, A, Rubin, D. *Bayesian Data Analysis*. Third. Chapman & Hall/CRC; 2013.
- Gelman A, Rubin DB. 1992; Inference from iterative simulation using multiple sequences. *Statistical Science*. 27:457–511.
- Gottman J, Swanson C, Swanson K. 2002; A general systems theory of marriage: nonlinear difference equation modeling of marital interaction. *Personality and Social Psychology Review*. 6:326–340.
- Hogan JW, Roy J, Korkontzelou C. 2004; Biostatistics tutorial: Handling dropout in longitudinal data. *Statistics in Medicine*. 23:1455–1497. [PubMed: 15116353]
- Hjort, N, Holmes, C, Muller, P, Walker, SG, editors. *Bayesian nonparametrics: principles and practice*. Cambridge University Press; 2010.
- Hogan JW, Laird NM. 1997; Model-based approaches to analysing incomplete longitudinal and failure time data. *Statistics in Medicine*. 16:259–272. [PubMed: 9004396]
- Ibrahim JG, Molenberghs G. 2009; Missing data methods in longitudinal studies: a review. *Test*. 18:1–43. [PubMed: 21218187]
- Lange KJ. 1989; Robust statistical modeling using the *t* distribution. *Journal of the American Statistical Association*. 84:881–896.
- Little RJA. 1993; Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association*. 88:123–134.
- Little RJA. 1994; A class of pattern-mixture models for normal missing data. *Biometrika*. 81:471–483.
- Little RJA. 1995; Modeling the drop-out mechanism in repeated measures studies. *Journal of the American Statistical Association*. 90:1112–1121.
- Little, RJA. Modeling the drop-out mechanism in repeated measures studies. In: Fitzmaurice, G, Davidian, M, Verbeke, G, Molenberghs, G, editors. Chapter 18 in *Advances in Longitudinal Data Analysis*. London: CRC Press; 2008.
- Little RJA, et al. 2012; The Prevention and Treatment of Missing Data in Clinical Trials. *N Engl J Med*. 367:1355–1360. [PubMed: 23034025]
- Michiels B, Molenberghs G, Lipsitz SR. 1999; Selection models and pattern-mixture models for incomplete data with covariates. *Biometrics*. 88:978–983.
- Rotnitzky A, Robins JM, Scharfstein DO. 1998; Semiparametric regression for repeated outcomes with non-ignorable non-response. *Journal of the American Statistical Association*. 93:1321–1339.
- Roy J, Daniels MJ. 2008; A general class of pattern mixture models for non ignorable dropout with many possible dropout times. *Biometrics*. 64:538–545. [PubMed: 17900312]
- Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. 2002; Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*. 64:583–639.
- Verbeke, G, Molenberghs, G. *Linear Mixed Models for Longitudinal Data*. New York: Springer-Verlag; 2000.
- Wu MC, Carroll RJ. 1988; Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*. 44:175–188.
- Wu MC, Bailey KR. 1989; Estimation and comparison of changes in the presence of informative right censoring: Conditional linear model. *Biometrics*. 44:939–955.
- Yuan Y, Little RJ. 2009; Mixed-effect hybrid models for longitudinal data with non-ignorable dropout. *Biometrics*. 65:478–486. [PubMed: 18759842]
- Zhang G, Yuan Y. 2012; Bayesian modeling longitudinal dyadic data with non ignorable dropout, with application to a breast cancer study. *The Annals of Applied Statistics*. 6:753–771. [PubMed: 23814631]

Zich JM, Attkisson CC, Greenfield TK. 1990; Screening for depression in primary care clinics: the CES-D and the BDI. *Int J Psychiatry Med.* 20:259–277. [PubMed: 2265888]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Sensitivity analysis of parameter estimates when the values of $\zeta_{12}^{(D)} = \zeta_{22}^{(D)}$ vary from -2 to 2 .

Table 1

The number of missingness classified by patients and spouses at different visit times.

	Dropout times		
	Baseline	3 months	6 months
Patient	0	108(27) [‡]	97
Spouse	0	108(25) [‡]	98

[‡]The numbers in parenthesis indicate the number of intermittent missingness cases.

Table 2

Simulation results, including the relative bias (RB=bias/truth), mean squared error (MSE), and coverage probability (CP) of the 95% credible (or confidence) interval for parameter estimates under the dyadic shared-parameter model (DSPM), the dyadic generalized linear mixed model (DGMM), the dyadic generalized linear mixed model before deletion (DGMM-BD), and the generalized linear mixed model before deletion (GMM-BD) under ignorable (Ignor) or non-ignorable (Nonign) missing data mechanisms.

Sc.	Mechanism	Parameters			Parameters		
	Dropout	Intermittent	β_1	$\tilde{\beta}_1$	β_2	$\tilde{\beta}_2$	$\tilde{\beta}_2$
	Before Deletion		GMM-BD			DGMM-BD	
		RB	-.089	.076	-.083	.090	.004
		MSE	.036	.007	.031	.009	.004
		CP	.177	.554	.245	.468	.946
1	Ignor		DGMM			DSPM	
		RB	0.029	-0.04	0.031	-0.039	0.008
		MSE	0.023	0.01	0.029	0.011	0.020
		CP	0.935	0.94	0.932	0.929	0.931
2	Nonign		RB	0.047	-0.062	0.043	-0.061
		MSE	0.038	0.015	0.039	0.014	0.019
		CP	0.891	0.923	0.892	0.898	0.936
3	Nonign		RB	0.033	-0.055	0.035	-0.052
		MSE	0.030	0.013	0.032	0.012	0.021
		CP	0.903	0.922	0.907	0.924	0.918
4	Nonign		RB	0.059	-0.043	0.041	-0.068
		MSE	0.050	0.014	0.039	0.017	0.026
		CP	0.857	0.931	0.891	0.894	0.913
5	Nonign		RB	0.068	-0.053	0.041	-0.067
		MSE	0.052	0.014	0.036	0.015	0.026
		CP	0.849	0.932	0.912	0.898	0.925

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Sc.	Mechanism		Parameters				Parameters			
	Dropout	Intermittent	β_1	$\tilde{\beta}_1$	β_2	$\tilde{\beta}_2$	β_1	$\tilde{\beta}_1$	β_2	$\tilde{\beta}_2$
6	Nonign	Nonign	RB	0.045	-0.06	0.058	-0.058	0.007	-0.026	0.018
			MSE	0.041	0.015	0.045	0.014	0.025	0.01	0.025
			CP	0.878	0.914	0.867	0.916	0.912	0.944	0.931
7	Nonign (misspecified)	Nonign	RB	0.041	-0.049	0.045	-0.056	0.003	-0.011	0.001
			MSE	0.035	0.013	0.043	0.014	0.02	0.009	0.023
			CP	0.894	0.919	0.867	0.905	0.925	0.941	0.908
8	Nonign (misspecified)	Nonign	RB	0.048	-0.055	0.048	-0.063	0.011	-0.016	0.007
			MSE	0.039	0.014	0.042	0.015	0.02	0.008	0.022
			CP	0.888	0.916	0.866	0.889	0.941	0.952	0.925

Parameter estimates and 95% highest posterior density (HPD) credible intervals (shown in parentheses) for the transition model.

Table 3

Missingness	Patients		Spouses	
	Parameter	Estimates	Parameter	Estimates
Intermittent	MPI	$\eta_{11}^{(\mathcal{J})}$	$\eta_{21}^{(\mathcal{J})}$	-1.05 (-1.42, -0.72)
	Random Slope	$\zeta_{11}^{(\mathcal{J})}$	$\zeta_{21}^{(\mathcal{J})}$	-0.17 (-0.84, 0.45)
Dropout	MPI	$\eta_{11}^{(\mathcal{D})}$	$\eta_{21}^{(\mathcal{D})}$	-0.28 (-0.5, -0.08)
	Random Slope	$\zeta_{11}^{(\mathcal{D})}$	$\zeta_{21}^{(\mathcal{D})}$	-0.07 (-0.61, 0.48)

Table 4

Parameter estimates and 95% HPD credible intervals (shown in parentheses) for the outcome process under the proposed DSPM and the dyadic generalized linear mixed model (DGMM).

		DSPM	DGMM
Patient	Intercept 1	1.38 (0.90, 1.87)	1.39 (0.90, 1.86)
	Intercept 2	2.98 (2.40, 3.58)	2.97 (2.38, 3.52)
	Time	−0.33 (−0.65, −0.04)	−0.43 (−0.72, −0.13)
	Patient MPI	0.41 (0.23, 0.59)	0.41 (0.23, 0.58)
	Spouse MPI	0.15 (0.01, 0.30)	0.16 (0.02, 0.32)
Spouse	Intercept 1	1.63 (1.13, 2.12)	1.58 (1.06, 2.05)
	Intercept 2	3.12 (2.55, 3.71)	3.12 (2.50, 3.68)
	Time	−0.16 (−0.47, 0.11)	−0.46 (−0.79, −0.13)
	Patient MPI	0.47 (0.3, 0.67)	0.45 (0.26, 0.61)
	Spouse MPI	−0.08 (−0.26, 0.09)	−0.07 (−0.24, 0.11)

Table 5

Parameter estimates and 95% HPD credible intervals (shown in parentheses) for the random-effects covariance under the proposed DSPM and DGMM.

Parameter	DSPM	DGMM
ρ_6	0.12 (0.02, 0.22)	0.05 (−0.07, 0.16)
ν_{11}	0.92 (0.83, 1)	0.87 (0.77, 0.96)
ν_{21}	1.08 (0.98, 1.17)	1.17 (1.06, 1.30)